NCCT Project Descriptions
BOSC Review, December 17-18 2007

Project Title: ToxCast™ - Developing Predictive Signatures for Chemical Toxicity

Lead Investigator: David Dix

NCCT Participants: Elaine Cohen Hubal, Rocky Goldsmith, Keith Houck, Richard Judson, Robert Kavlock, Stephen Little, Matt Martin, Keith McLaurin, James Rabinowitz, David Reif, Ann Richard, Daniel Rotroff, Woodrow Setzer.

1. Short description of the topic/project:

There is a clear international need to develop predictive tools for evaluating the potential of chemicals to induce toxicity and the mechanisms by which they do so. For example, it has been estimated that adequate toxicology information is available for less than 20% of the chemicals of concern to the US EPA (Judson et al 2008). Having scientifically accepted predictive tools will enable the more efficient and effective characterization of chemical risk and hazard, and lead to a more rational use of animals in research as they are directed to the highest priority chemicals. Historically, most attempts to meet the need for predictive tools have focused on the development of quantitative structure-activity relationship (QSAR) models. While QSAR models have proven useful for predicting some mechanisms within relatively well defined chemical classes, they often do not perform well when screening broadly diverse categories of chemicals because their training sets do not adequately cover the relevant chemical space. Therefore other approaches are needed, and the experience of the pharmaceutical industry in the use of state-of-the-art high-throughput screening assays (HTS), toxicogenomics, and computational chemistry tools for the discovery of new drugs might provide a solution. HTS refers to a system that rapidly and efficiently tests large numbers (i.e., thousands) of chemicals for bioactivity, typically utilizing robotics and automation applied to biochemical and cellular assays (Inglese et al 2006). Exploiting recent advances in HTS and toxicogenomics, EPA has launched a research program called ToxCast™ to develop methods for prioritizing chemicals for further screening and testing (Dix et al 2007).

2. What is the EPA context for the project?

The National Research Council (NRC) of the National Academy of Sciences (NAS) provided EPA a report in June 2007 entitled "Toxicity Testing in the Twenty-first Century: A Vision and a Strategy" (NRC, 2007). The report's overall objective is to foster a transformative paradigm shift in toxicology based largely on the use of in vitro systems that will (1) provide broad coverage of chemicals, chemical mixtures, outcomes, and lifestages, (2) reduce the cost and time of testing, (3) use fewer animals and cause minimal suffering in the animals used, and (4) develop a more robust scientific base for assessing health effects of environmental agents. This vision is highly consistent with the strategic directions and research activities of EPA. Even before the EPA commissioned this report, it was taking steps to incorporate modern biological and computational tools into the evaluation of chemicals hazards and risks. ToxCast is EPA's major effort to predict hazard, identify key toxicity pathways, and intelligently prioritize chemicals for targeted, hypothesis-driven animal testing. The endorsement of this approach by the NRC provides further assurance that the Agency's ToxCast program is on the right track. Developing ToxCast within EPA, an organization that uses cutting

edge tools at the bench level, and is responsible for protecting human health and the environment, will facilitate regulatory acceptance as the science unfolds. The NCCT is working across the Agency, with other Federal agencies, and with stakeholder groups to develop this help advance the transformation called for by the NRC.

ToxCast is a departure from mainstream toxicology, and will require regulatory acceptance by the Agency and stakeholders alike for successful applications in chemical prioritization. Unlike previous single assay efforts to develop alternative assays and their validation through procedures utilized by ICCVAM and ECVAM, ToxCast poses some unique challenges due to its multidimensional nature and use of assays available only from single sources. Experience with Phase I and II of the program should provide significant guidance for the development of a scientific consensus as to the validity and acceptability of the approach. The transparent nature of the program, with its full public release of data, has been a guiding principle in order to facilitate scientific acceptance.

3. What are the strategic directions and science challenges?

The pharmaceutical industry, in its efforts to develop more efficient methods for drug discovery, is responsible for many of the new tools available to toxicology. While overall assessments of the success of that industry are cloaked in confidential business information, there have been enough glimpses of their potential to detect toxicity pathways to show that this should be a promising approach to be applied to environmental chemicals (Houck and Kavlock, 2007). Nevertheless, there are numerous differences and challenges in translating the experience of drug development to environmental toxicity. Notable of these are: (1) pharmaceutical compounds are developed to be biologically active, and hence may be more amenable to such an approach than environmental chemicals, many of which have no intended biological activity; (2) the chemical space covered by drug development is considerably narrower than that of environmental chemicals, which do not have to display particular ADME characteristics that make drugs bioavailable and efficacious; (3) the metabolism of drugs is generally well characterized, as is the activity of any metabolites; (4) selection of chemical libraries used for drug discovery include consideration of solubility, which is not necessarily of importance for environmental chemicals; (5) while there may be only 400-500 drugable therapeutic targets, the number of potential toxicological targets for drugs and environment chemicals is likely to be quite large, necessitating a very broad search for toxicity pathways; (6) for assessment of environmental chemicals, confidence in the value of a negative result in any assay will have to be higher than for drugs, which will certainly undergo additional scrutiny as preclinical studies continue on lead compounds; (7) the probability is high that environmental chemicals are going to interact with more than a single toxicity pathway, with exposure intensity and duration playing important determinants as to which one(s) are key to inducing adverse health effects; (8) the likelihood that the outcomes of perturbing toxicity pathways will be cell-type dependent, reflecting the variability in expression of specific molecular targets; and (9) the aspect that at least some forms of toxicity are dependent on higher order interactions of cells in tissues or organs what may not be apparent by a reductionist evaluation of isolated cells and/or pathways. Even this limited list of obstacles exposes the daunting challenge in creating a new paradigm for toxicological evaluation, and it also points to the need for a strategic approach to design a research program that can begin to chip away at the obstacles.

The application of HTS in toxicology could be twofold, one essentially bottom up, the second top down. In the bottom up approach, a single or small number of chemicals could be analyzed against a vast array of targets to isolate the key toxicity pathways. In the top down approach, a relatively large number of chemicals could be assayed against

a small number of key targets. This is essentially the approach being used for detection of endocrine disrupting chemicals that act via interaction with estrogen, androgen or thyroid hormone function. A middle ground would try to maximize both the numbers of chemicals assayed and the breadth of the assays so as to achieve a biologically based prioritization process. ToxCast is designed to test the hypothesis that multi-dimensional evaluation of chemical properties and effects across a broad spectrum of information domains (e.g., molecular, cellular, and organ responses) will provide data that will be predictive of toxicity. The goal is to acquire sufficient information on a range of chemicals, so that "bioactivity signatures" can be discerned that identify distinctive patterns of toxic effects, or phenotypes, observed in traditional animal toxicity testing. The ToxCast predictive bioactivity signatures will be based upon physical-chemical properties, predicted biological activities from structure-activity models, biochemical properties from HTS assays, cell-based phenotypic assays, genomic analyses of cells in vitro, and responses in non-mammalian model organisms. The ToxCast assays will provide information, quickly and in a cost-efficient manner, on the potential impact of chemicals on numerous biological pathways critical for the function of systems such as the heart, lungs, liver, brain or reproductive organs.

4. What are the short-term (1-2 year) and long-term (3-5 year) goals?

This five-year effort is divided into three phases (Table 1). In Phase I, as a proof-of-concept, ToxCast is examining more than 300 chemicals, in over 400 different HTS bioassays, to create predictive bioactivity signatures at a cost of less the $20,000 per chemical. The Phase I chemicals are primarily pesticide active ingredients that have been extensively evaluated by traditional mammalian toxicity testing, and hence have known properties representative of a number of phenotypic outcomes (e.g., carcinogenicity; and developmental, reproductive and neural toxicity). A $6M investment has been made to date in Phase I. A large number of Phase I results are already in hand, and completion of data generation is anticipated by the spring of 2008. Analysis and an initial set of ToxCast predictive signatures are expected by summer of 2008, and the subsequent Phase II will focus on the confirmation and expansion of these predictive signatures by generating additional HTS data on up to 1000 more chemicals. In Phase III, ToxCast will be expanded to the thousands of environmental chemicals, delivering an affordable, science-based system for categorizing chemicals. As the ToxCast database grows, so will confidence in predicted toxicity and potential mechanisms of action useful in refining and reducing the use of animals in toxicity testing. To ensure transparency and collaboration, ToxCast data will be freely available on the internet.

5. What other components of EPA or outside organizations are involved?

To advance research into the utility of computational chemistry, HTS and various toxicogenomic technologies for Agency use and broaden awareness and participation in this area, the EPA Chemical Prioritization Community of Practice (CPCP) was formed in December 2005. The goal of the CPCP is to provide a venue for stakeholder information sharing and discussion related to these technologies and for interpretation in order to categorize chemicals and predict toxicity. The CPCP is chaired by the NCCT and meets monthly by teleconference. It has a membership of over 100 individuals from 20 public and private sector organizations.

Internationally, the Organization of Economic Cooperation and Development (OECD) has supported a project proposal developed jointly by the NCCT and the Office of Pollution Prevention and Toxic Substances (OPPTS) to promote international cooperation and research on application of new molecular based approaches for the prioritization and screening of environmental chemicals for potential toxicity. The

objective of the "Molecular Screening for Characterizing Individual Chemicals and Chemical Categories Project" project is to establish a strategy for rationally and economically prioritizing chemicals for further evaluation, based on molecular properties and categories linked to potential toxicity. This objective directly builds on the goals of the ToxCast program, and the needs of EPA relevant to various chemical programs. Recognizing the need for international acceptance and harmonization of molecular screening tools, the NCCT and OPPTS approached OECD about facilitating such an activity. The project was formally accepted by the joint OECD/International Programme on Chemical Safety Advisory Group on Toxicogenomics in January 2007, and a workshop was held in May 2007 to initiate collaborative efforts. It is likely that several countries and companies will become active participants in the effort. In the 2008 – 2009 timeframe, further development of partnering arrangements, infrastructure, and information sharing will occur. Also during this timeframe, a specific list of chemicals and methodologies should be agreed upon for the OECD-coordinated effort. The OECD Molecular Screening Project represents a valuable opportunity for the Agency to link the ToxCast program to international research trying to develop solutions to the increasing demand for chemical testing, and provide science-based prioritization for the toxicity testing of environmental chemicals.

6. How is data management being achieved?

The potential for ToxCast to use HTS and genomic data for environmental chemical hazard assessment, screening, and prioritization requires initial anchoring of these data to reference toxicological test information. For this end, ToxRefDB (Toxicology Reference Database) is being created to provide a relational database of standard toxicity test results for pesticides, making it possible to link toxicity information with the HTS and genomic data of ToxCast. ToxRefDB uses a controlled ontology and standardized data field structure to capture toxicological endpoints, critical effects and relevant dose-response data (Martin et al 2008). The creation and population of ToxRefDB has been a collaborative effort between EPA Office of Research and Development National Center for Computational Toxicology (ORD/NCCT) and EPAs Office of Pesticide Programs (OPP). ToxRefDB currently provides the ability to cluster and group chemicals based on toxicological outcomes specific to study type, target organ, or effect categories. In addition, ToxRefDB will facilitate ranking of chemicals by relative potency based on specific endpoints, or grouping of chemicals based on mode or mechanism of action. Thus, ToxRefDB will provide the essential interpretive context for linking ToxCast HTS and genomic data to toxicity endpoints. It has been designed to be scaled up to capture standard toxicity information for chemicals that will be part of Phase II.

In order to integrate the relational environment of ToxRefDB with associated chemical structure information, and tools for searching and categorization that will guide development of predictive HTS bioactivity profiles and genomic signatures, a more comprehensive data management system is being developed by NCCT. ACToR (Aggregated Computational Toxicology Resource) will manage the large-scale sets of ToxCast assay data, and is comprised of several independent data repositories, tied together through links to a common database of chemical structures and properties. The main databases cover chemical information, biochemical (HTS) and cell-based assays, detailed in vivo toxicology data (ToxRefDB) , experimental design information, genomics data (mainly microarray), and reference information on genes and pathways. ACToR is collecting information from multiple sources both within and external to the EPA. Users will be able to access data through the EPA internet (www.epa.gov/ncct/toxcast).

7. What are appropriate measures of success?

Bioactivity signatures will be defined and evaluated by their ability to predict outcomes from existing mammalian toxicity testing. These hazard predictions should provide the public, the chemical industry and governmental programs around the world with science-based information helpful in prioritizing chemicals for more detailed toxicological evaluations, and therefore lead to using animal tests more efficiently. The results will provide, for the first time, a comprehensive and detailed overview of the potential impact of environmental chemicals upon key cellular activities. The assays range from characterizing the interactions of chemicals with proteins that regulate and maintain proper cell function, to measuring the response of whole cells, to studying chemical effects in a model organism (Table 2),

ToxCast, with its multi-dimensional approach and supporting informatics infrastructure on traditional toxicity testing results, holds the promise to complement and expand existing chemical screening approaches by efficiently and quantitatively prioritizing EPA-relevant chemicals based on computational models using chemical descriptors and biological activity profiling in Phases II and III (Table 1). Armed with this science-based information, EPA programs can further prioritize chemicals for more detailed evaluations, including using animal tests more efficiently. The EPA is very interested in continuing to engage other organizations in collaborative research arrangements in which we share experiences, efforts, and best practices relevant to ToxCast, HTS screening, and chemical prioritization. Several examples of key collaborations within EPA, with stakeholders, and with international partners will be discussed.

8. References

Dix, D.J., Houck, K.A., Martin, M.T., Richard, A.M., Setzer, R.W., Kavlock, R.J., 2007. The ToxCast Program for Prioritizing Toxicity Testing of Environmental Chemicals. *Toxicol. Sci*. 95, 5-12.

Houck, K.A. and Kavlock, R.J. (2007). Understanding mechanisms of toxicity: Insights from drug discovery. *Toxicol and Appl. Pharm.* (in press).

Inglese, J., Auld, D.S., Jadhav, A., Johnson, R.L., Simeonov, A., Yasgar, A., Zheng, W. and Austin, C.P. (2006). Quantitative high-throughput screening: a titration-based approach that efficiently identifies biological activities in large chemical libraries. *Proc. Natl. Acad. Sci. USA* **103**, 11473-11478.

Martin MT, Houck KA, McLaurin K, Richard AM, Dix DJ (2007). Linking Regulatory Toxicological Information on Environmental Chemicals with High-Throughput Screening (HTS) and Genomic Data. *The Toxicologist CD- An official Journal of the Society of Toxicology*, **96:**219-220.

National Research Council (2007). Toxicity Testing in the 21st Century: A Vision and a Strategy. National Academies Press, Washington DC.

Table 1. Phased Approach to Development of ToxCast Signatures

Phase	Number of Chemicals	Chemical Criteria	Purpose	Est. Cost per Chemical	Target Date
I	>300	Data Rich (pesticides)	Signature Development	$20k	FY07-08
II	>1000	Expanded Structure and Use Diversity	Evaluation and Extension	$15-20k	FY08-09
III	Thousands	Data poor	Prediction and Prioritization	$10-15k	FY10-12

Table 2. Endpoints contained within Phase I signature development					
Assay Type	Number of Assays	Number of Unique Endpoints	Assay Source	Comment	Source
Biochemical	240	240	Mostly human and rat	Enzyme inhibition, Ion channels, GPCRs, Cytochromes	NovaScreen Biosciences
Transcription Factor Profiling	2	67	HepG2 cells (human liver)	Nuclear receptors and other transcription factors	Attagene
Nuclear receptor activation	10	10	Human and rodent	Reporter gene assay over 15 concentrations	NIH Chemical Genomics Center
Transcriptomics	1	22,000	Primary hepatocytes-Kupffer cell co-cultures	Illumina microrrays	In Vitro ADMET Laboratories and Expression Analysis
Kinetic Cell Growth	1	Kinetic	A549 cells (human lung)	Real time recording of electrical impedance	ACEA Biosciences
Cytotoxicity and Bioactivation	1	6	Primary human liver, lung and kidney cells	Shared metabolism across cell types	In Vitro ADMET Laboratories
Complex cell culture	8	87	Primary human cells	Many cell signaling pathways	Bioseek
High content screening	1	11	HepG2 cells (human liver)	Fluorescence imaging of cells	Cellumen
Fish development	1	11	Zebrafish (Dana rerio)	Teratogenesis	Phylonix
TOTAL	265	22,433			